Artificial Neural Networks for SCADA Data based Load Reconstruction

Claudia Hofemann, Gerard van Bussel, Herman Veldkamp
DUWIND, Delft University of Technology, the Netherlands

ABSTRACT
If at least one reference wind turbine is available, which provides sufficient information about the wind turbine loads, the loads acting on the neighbouring wind turbines can be predicted via an artificial neural network (ANN). This research explores the possibilities to apply such a network not only within a wind park but on turbines located at different sites. Following the idea to develop a tool to forecast the particular loads of any wind turbine in the field without the need to install additional measuring systems, a model has been developed needing only signals contained in the Supervisory Control and Data Acquisition (SCADA) data as input signals.

INTRODUCTION
This study is part of a project investigating the potential of data provided by the SCADA system to forecast the site dependent mechanical loading of wind turbines over their lifetime. Standard signals, providing online information, are available for nearly all commercial wind turbines. Those signals, obtained by the SCADA system, provide turbine operators and manufactures with online information of the wind turbine state, the state of turbine components and the environmental conditions such as temperature or wind speed. Artificial neural networks (ANN) have been successfully applied to estimate the loads acting on a particular wind turbine [1] or to assess the performance of a wind farm [3] on the condition that verification data is available. Here as well an ANN has been trained, but this time the network is not only tested on the capability to reconstruct the loads but also on its transferability. Bringing us a step closer to the underlying idea of those projects: to establish a system allowing to estimate the loads without the need to install additional measurement equipment. Up to now the developed networks (as described in [1] and [3]) have been only tested on one particular turbine or within one Wind Park. Here ANN is trained for one turbine and tested on a different turbine located at a different site. The by the ANN established empirical relationships are based on SCADA data as input data and tested by verification data obtained from load measurements. The verification data represents the actual loads measured on the turbine. Verification data is only needed to validate the code. Once the ANN is trained, predictions are performed using only the input data; hence signals which are included in the standard SCADA data. Implying that the wind turbine loads can be estimated without the need to install additional load sensors.

ARTIFICIAL NEURAL NETWORKS
ANNs can be used for a wide range of applications. Up to now ANN’s are not that common in the wind turbine industry but have been applied successfully on other forecasting problems as the development of the stock market or the energy price. From an engineering point of view most of the ANNs can be seen as a black box. They are inspired by the mechanism of the brain and can be classified by different categories as for example their learning mechanism or how they are trained (supervised \ unsupervised) [2]. ANN’s accept input parameters, process them, and produce output parameters according to a nonlinear transfer function [4]. For the described forecasting problem a recurrent feed forward neural network, a Nonlinear Autoregressive Exogenous Model (NARX) has been chosen. Those network types are typically applied to problems dealing with time series predictions.

TRAINING OF THE ANN
The network is trained to predict the blade root moments of two wind turbines located at different sites. A supervised ANN has been chosen, implementing that load data has to be provided to train the network to establish the relationship between the SCADA data and the measured loads. The load data is the target data of the network and used to train the network as well as to verify the made predictions. Parameters as wind speed, wind direction, power output, pitch angle and yaw moment are used as input data. The predictions are performed by a dynamic recurrent feed forward neural network trained with the Bayesian regulation back
propagation, a network training function that updates the weight and bias values according to Levenberg-Marquardt optimization. It minimizes a combination of squared errors and weights, and then determines the correct combination to produce a network that generalizes well [2]. The network has been trained with 10 minutes statistics. The period of 10 minutes has been chosen for the simple reason that the statistics transmitted by the SCADA system, are usual parameters computed over a period of 10 minutes. The data which has been used to train the network is not filtered or sorted and chronologically presented to the network.

PERFORMANCE OF THE MODEL
The fully developed model shall perform predictions of the wind turbine loads of all wind turbines of at least one fleet. The model has been tested on three different turbines. The results regarding the reconstruction of the loads on one particular turbine are promising. The transferability of the model has been tested with two wind turbines, both located in a similar environment: the reference wind turbine, needed to train and verify the model and a second wind turbine to test the developed model.

Seven load sensors are installed on the blades of the reference wind turbine. The model is trained with all of the available sensors and tested with the available set from the test turbine, taking in account small inaccuracies due to a different mounting of the sensors. The signals form missing sensors are set to zero. From the available set of signals representing the SCADA data, 18 different input parameters as wind speed, wind direction and turbine state have been chosen to train the neural network.

![Figure 1](image.png)

Figure 1. Normalised blade bending moments plotted over the time for two different wind turbines. For each of the turbines the simulated output form the network is given as well as the original target data.

Figure 1 illustrates four time series, generated for the two different turbines. The simple lines; represent the measured blade bending moments, the target data of the network. The dashed lines represent the simulated output from the model. The reconstruction of the blade bending moments for one particular wind turbine is accurate as shown by the red and the orange lines of figure 1. If the model has been trained with sufficient data gives a correlation coefficient between $R=0.95$ and $R=0.99$.

Transferring the established relationships from the reference turbine to the test turbine, by applying new data from the test turbine to the original network, gives a correlation coefficient of $R=0.93$. For this case no
additional training has been performed with the data from the test turbine. This case is visualized by the blue and green lines in figure 1. If additional training is performed the correlation coefficient increases to $R=0.97$.

OUTLOOK
The established knowledge structures can be transferred from one turbine to another, proving that it is possible to reconstruct the wind turbine loads within one fleet utilizing only SCADA data. The achieved accuracy of the model is promising but not satisfying. Further modifications of the model are needed to increase the reconstruction capabilities. Until now tests have been performed with two turbines located in a similar environment, in the future the model shall be tested under different environmental conditions. A test of the model in complex terrain as well as offshore is planned for the near future.

ACKNOWLEDGMENTS
I would like to thank Vestas for supporting this project, in particular Srikar Vulli and Karl H. Svendsen.

BIBLIOGRAPHY